# SUPERVISOR'S OPINION OF FINAL THESIS

**I. IDENTIFICATION DATA**

| | |
|---|---|
| **Thesis name:** | **Relaxed Quantization and Binarization for Neural Networks** |
| **Author's name:** | **Mráz Martin** |
| **Type of thesis :** | bachelor |
| **Faculty/Institute:** | Faculty of Electrical Engineering (FEE) |
| **Department:** | Department of Cybernetics |
| **Thesis supervisor:** | Shekhovtsov Oleksandr Mgr.,Ph.D. |
| **Supervisor's department:** | Department of Cybernetics |

**II. EVALUATION OF INDIVIDUAL CRITERIA**

**Assignment** **challenging**

*Evaluation of thesis difficulty of assignment.*

The assignment required skills in programming, deep learning, probability theory and statistics. It consisted in studying a rather broad spectrum of existing methods for the problem, experimental comparison, individually deriving a probabilistic propagation method, deploying it to a realistic application and exploring further directions of improvements. The complexity was balanced by organizing the work as a collaboration. The thesis explicitly states what parts were implemented independently and what were implemented within a software framework developed by the advisor.

**Satisfaction of assignment** **fulfilled with minor objections**

*Assess that handed thesis meets assignment. Present points of assignment that fell short or were extended. Try to assess importance, impact or cause of each shortcoming.*

The thesis builds on the preceding semestral project of Martin. It was thus possible to accomplish main steps, which was challenging from booth the theoretical and software engineering perspectives. The main goals of comparing estimators, deriving and implementing probabilistic propagation methods were achieved. Experiments on small datasets did not confirm expected speed-up, they however revealed important information for rethinking the method in the future. Focusing on the probabilistic propagation itself and having little success with it in practice so far, we ran out of time to reach more distant optimistic goals: using it in large-scale training for realistic applications, more broad comparison, refining how a single deterministic model is obtained. This is not to be considered negatively.

**Activity and independence when creating final thesis** **B - very good.**

*Assess that student had positive approach, time limits were met, conception was regularly consulted and was well prepared for consultations. Assess student's ability to work independently.*

Martin was actively catching up with deep learning in general and the literature in the scope of the thesis, systematic in organizing his notes and writing down the material. The time limits on achieving particular steps were not set by me explicitly, rather I posed sub-tasks and then we iterated on them as long as it was necessary. It was visible that Martin is putting efforts into it and each time progressed by either learning or deriving or implementing some steps. It took more time than I initially expected and required substantial guidance, however I can conclude that Martin learned efficiently from the beginning of the project. He completed steps of independent derivation of probabilistic propagation and independent implementation tasks and has taken some steps beyond, which was a useful exercise. The thesis was written mostly independently, with just few general advices and corrections from me.

**Technical level** **B - very good.**

*Assess level of thesis specialty, use of knowledge gained by study and by expert literature, use of sources and data gained by experience.*

Martin surveyed stochastic relaxation approach to trained quantization, sampling based gradient estimators and derived estimator based on approximating the expectation analytically by propagating means and variances of activations distributions. All these parts required understanding of technically demanding literature, working formally with statistical

tools and implementing methods using non-standard approaches in pytorch.

**Formal and language level, scope of thesis** **C - good.**

*Assess correctness of usage of formal notation. Assess typographical and language arrangement of thesis.*

There are some issues with notations and equations, mainly in Part 2. Some explanations are not clear enough and some are in fact misleading. I believe Martin has learned and understood all the methods involved, but perhaps not to the level to present them clearly and concisely and in a clear relation with the problem and to each other. Sometimes this is due to unclear English. For some parts, technical details could be more complete. All of this is of course is difficult to achieve at a bachelor level.

**Selection of sources, citation correctness** **B - very good.**

*Present your opinion to student's activity when obtaining and using study materials for thesis creation. Characterize selection of sources. Assess that student used all relevant sources. Verify that all used elements are correctly distinguished from own results and thoughts. Assess that citation ethics has not been breached and that all bibliographic citations are complete and in accordance with citation convention and standards.*

Martin red and carefully annotated all recommended papers, which we subsequently discussed. Some of them appeared to be a bit too difficult though. Nevertheless he got a good overview and sufficient understanding to work with all methods presented in the thesis. All used elements from the literature and those provided by the advisor are correctly distinguished from own results and thoughts. Citation ethics has not been breached and that all bibliographic citations are complete and in accordance with citation convention and standards. Martin has independently explored and cited in the thesis some research papers I was not aware of. To criticize a bit, the binary special case, and in particular the recommended reference [1] from the assignment, are not covered. In particular STE for the stochastic relaxation is derived there as a theoretically sound method, in contrast to some claims in the conclusion of the thesis.

**Additional commentary and evaluation**

*Present your opinion to achieved primary goals of thesis, e.g. level of theoretical results, level and functionality of technical or software conception, publication performance, experimental dexterity etc.*

To summarize, Martin has learned stochastic quantization methods, an approach which is based on theoretical understanding (unlike many quantization heuristics in the field). The experimental evidence corresponding to part 2 confirmed that there are alternatives to relaxed quantization, which are much simpler and are as efficient and theoretically sound. In part 3, Martin has independently derived probabilistic propagation method in order to gain understanding and extend it. The use of the derived method for the proposed speed-up purpose did not fulfill practical expectations yet but revealed some issues to check in the future.

Overall, I am positive about the progress of Martin in learning and gaining qualification on a challenging level. However I must honestly say, that his contribution to either experiments or theory does not allow speaking of co-authorship of any possible future scientific publications on the topic.

ČVUT
ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

# SUPERVISOR'S OPINION OF FINAL THESIS

**III. OVERALL EVALUATION, QUESTIONS FOR DEFENSE, CLASSIFICATION SUGGESTION**

The extend of the thesis covers two families of methods, relatively well presented and illustrated, of challenging technical difficulty. The work is complete in the sense that it contains all parts: motivation, survey of methods, an independently carried theoretical derivation (part 3 + appendix); the methods are implemented and compared.
Questions for defense:

1. Please correct the formal explanation of stochastic rounding. In eq. (2.8) two different uniform distributions are specified (the equality is false) -- which one is correct? Explain which \tilde x is continuous and which is discrete in (2.8) and Fig. 2.2. Why in Fig. 2.2, if x is closer to g_i, the probability to quantize it to g_i decreases (blue area), when it should increase? Can you derive the formula in the text below (2.8) from the noisy model (2.8) and the categorical distribution of noisy quantization (2.6)?
2. What would you do differently in the problem assignment, approaching the quantization problem or the work arrangement?

I evaluate handed thesis with classification grade **B - very good.**

Date: **30.5.2022** Signature:

ČVUT
ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

# POSUDEK OPONENTA ZÁVĚREČNÉ PRÁCE

## I. IDENTIFIKAČNÍ ÚDAJE

| | |
|---|---|
| **Název práce:** | **Relaxed quantization and Binarization for Neural Networks** |
| **Jméno autora:** | **Martin Mraz** |
| **Typ práce:** | bakalářská |
| **Fakulta/ústav:** | Fakulta elektrotechnická (FEL) |
| **Katedra/ústav:** | Katedra kybernetiky |
| **Oponent práce:** | Teymur Azayev |
| **Pracoviště oponenta práce:** | Katedra kybernetiky |

## II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

**Zadání** — náročnější

*The work consists of theoretical and programming work in a niche field which is an active area of research. This puts the difficulty to above average due to scarcity of materials and well established methods as well as the requirement of understanding new concepts which can only be found in scientific publications.*

**Splnění zadání** — splněno s většími výhradar

The first (main) point is fulfilled, the student implemented a framework for the proposed methods and compared the performance and shows that a simpler method (SR) is superior to more complex SQ method. Experiments show that probabilistic pretraining does not help, but the author suggests that it might, in larger networks.
Points 2 & 3 are not implemened. They seem to be stand-alone projects and would require a significant amount of time and work. Point 2 would require the student‘s creativity and experimentation to implement so it is a shame that it wasn‘t realised, but Point 3 is just an additional experiment to show performance so I don‘t think that it‘s important. In general, the supervisor likely underestimated the amount of time which would be required to implement all the points.

**Zvolený postup řešení** — správný

*The proposed solution is ok.*

A few minor comments:
It seems that mostly only training time and accuracy has been considered as the main criterium for the various approaches. It would be nice to see memory usage and runtime comparisons on several different systems (normal pc, single board computer (SBC), etc), since this is the main selling point of these quantization methods.

**Odborná úroveň** — B - velmi dobře

The technical level is very good. Explanations in intro and theoretical parts are usually comprehensive. Good illustrations and mathematical description. Figures could be better described in some cases.

ČVUT
ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

# POSUDEK OPONENTA ZÁVĚREČNÉ PRÁCE

**Formální a jazyková úroveň, rozsah práce** B - velmi dobře

*The formality and level of language is generally very good at the beginning, so it shows that the student is capable of it, but the experiments and discussion section seems a bit rushed, with some errors in language, sentence formulation and some lacking explanations.*

**Výběr zdrojů, korektnost citací** A - výborně

*The citations for the used and relevant work is comprehensive and satisfactory.*

**Další komentáře a hodnocení**

*A lot of what is in the discussion probably belongs in the intro. The experiments could be organised a bit better. Even though the proposed probabilistic pretraining didn't help, it would be nice to see the results in a table as well.*

## III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

*- Overall assessment: The work generally looks well done, with decent quality explanations, illustrations and mathematical derivations, however due to some un implemented points and rushed ending I would prefer to give the work a C (good) grade, but I suspect that the negative aspects of the work can be attributted to a mismanagement of the supervisor so I will give the work an official B (very good) grade.*

*- Question: For me the initial curiosity was not how to quantize a 1 or 2 bit network, but if such a network has the learning capacity to perform the trained task and how the degree of quantization has an effect on learning capacity in comparison to using high precision floats. Is there any way to quantify this or any work that looks at it?*

Předloženou závěrečnou práci hodnotím klasifikačním stupněm B - velmi dobře.

Datum: 1.6.2022 Podpis: